# Research

# Spatial Variation of Endotoxin Concentrations Measured in Ambient PM$_{10}$ in a Livestock-Dense Area: Implementation of a Land-Use Regression Approach

*Myrna M.T. de Rooij,[1] Dick J.J. Heederik,[1] Erik J.H.M. van Nunen,[1] Isabella J. van Schothorst,[1] Catharina B.M. Maassen,[2] Gerard Hoek,[1] and Inge M. Wouters[1]*

[1]Division of Environmental Epidemiology, Institute for Risk Assessment Sciences, Utrecht University, Utrecht, Netherlands
[2]National Institute for Public Health and the Environment (RIVM), Bilthoven, Netherlands

**BACKGROUND:** Results from studies on residential health effects of livestock farming are inconsistent, potentially due to simple exposure proxies used (e.g., livestock density). Accuracy of these proxies compared with measured exposure concentrations is unknown.

**OBJECTIVES:** We aimed to assess spatial variation of endotoxin in PM$_{10}$ (particulate matter $\leq 10$ μm) at residential level in a livestock-dense area, compare simple livestock exposure proxies to measured endotoxin concentrations, and evaluate whether land-use regression (LUR) can be used to explain spatial variation of endotoxin.

**METHODS:** The study area (3,000 km$^2$) was located in Netherlands. Ambient PM$_{10}$ was collected at 61 residential sites representing a variety of surrounding livestock-related characteristics. Three to four 2-wk averaged samples were collected at each site. A local reference site was used for temporal variation adjustment. Samples were analyzed for PM$_{10}$ mass by weighing and for endotoxin by using the limulus amebocyte lysate assay. Three LUR models were developed, first a model based on general livestock-related GIS predictors only, followed by models that also considered species-specific predictors and farm type–specific predictors.

**RESULTS:** Variation in concentrations measured between sites was substantial for endotoxin and more limited for PM$_{10}$ (coefficient of variation: 43%, 8%, respectively); spatial patterns differed considerably. Simple exposure proxies were associated with endotoxin concentrations although spatial variation explained was modest ($R^2 < 26\%$). LUR models using a combination of animal-specific livestock-related characteristics performed markedly better, with up to 64% explained spatial variation.

**CONCLUSION:** The considerable spatial variation of ambient endotoxin concentrations measured in a livestock-dense area can largely be explained by LUR modeling based on livestock-related characteristics. Application of endotoxin LUR models seems promising for residential exposure estimation within health studies. https://doi.org/10.1289/EHP2252

## Introduction

There is concern about the influence of air pollution from livestock farms on public health. Several studies have shown associations between livestock farming exposure proxies and respiratory health of neighboring residents (Borlée et al. 2015; Elliott et al. 2004; Mirabelli et al. 2006; Pavilonis et al. 2013b; Radon et al. 2007; Schiffman et al. 2005; Schinasi et al. 2011; Schulze et al. 2011; Sigurdarson and Kline 2006; Smit et al. 2014; Wing and Wolf 2000). Livestock farming is associated with emissions of gases and primary and secondary (formed from gases such as ammonia) particulates, including bioaerosols (Cambra-López et al. 2010; Hamon et al. 2012; Seedorf et al. 1998).

Bioaerosols are aerosolized dust particles originating from microbial, animal or plant materials. One very potent bioaerosol is endotoxin, an inflammatory component of the cell wall of Gram-negative bacteria (Liebers et al. 2008). Endotoxins are absorbed onto the surface of particles, mainly coarse particulate matter (Chen and Hildemann 2009; Morgenstern et al. 2006; Soukup and Becker 2001). Inhalation of increased endotoxin levels can cause respiratory effects and even systemic effects, resulting in a variety of symptoms (Basinas et al. 2015; Liebers et al.

2008). High endotoxin concentrations have been measured on farms (Clark et al. 1983; Jonges et al. 2015; Seedorf et al. 1998; Thorne et al. 2009), and negative effects on respiratory health were found among farmers (Basinas et al. 2015). Endotoxin is ventilated outward from farms using particulate matter as a vector (Pillai and Ricke 2002). Elevated endotoxin concentrations have been measured in the direct vicinity of farms; however, the concentrations measured outside the farms were much lower compared with levels measured inside (Jonges et al. 2015; Thorne et al. 2009). Still, the effect of livestock farming on endotoxin concentrations at the residential level is largely unknown because knowledge on emission and dispersion is scarce. Respiratory health effects observed in residents living near livestock farms have been suggested to be associated with endotoxin emitted from these farms (Heederik et al. 2007; Mirabelli et al. 2006; Schinasi et al. 2011; Schulze et al. 2006; Smit et al. 2017).

Different exposure proxies have been used to represent residential exposure to emissions from livestock farms in studies in populations living in the vicinity of livestock production units. Exposure proxies used included farm density of the region/county (Elliott et al. 2004; Sigurdarson and Kline 2006; Wing and Wolf 2000), weighted distance to nearest farm (Borlée et al. 2015; Mirabelli et al. 2006; Smit et al. 2014), and farm density in buffers of various sizes ranging between studies from 500 m to 5 km (Borlée et al. 2015; Pavilonis et al. 2013b; Radon et al. 2007; Smit et al. 2014). Some studies evaluated all farm types combined (Pavilonis et al. 2013b; Radon et al. 2007), others differentiated farm types by animal species (Borlée et al. 2015; Elliott et al. 2004; Smit et al. 2014; Wing and Wolf 2000) or focused on one particular animal species [pigs (Mirabelli et al. 2006; Sigurdarson and Kline 2006)]. In general, only farm numbers were taken into account, except for the study of Mirabelli et al. (2006), which also took into account herd size. Accuracy of exposure proxies in comparison with measured livestock associated exposures is unknown. This hampers interpretation of results and comparisons between epidemiological studies.

Knowledge of spatial variation, gained by performing measurements at a large number of sites, is essential to increase insight into residential exposure. Few studies have been performed that assess endotoxin concentrations in ambient air, and few of these took into account the relation between agriculture and ambient air levels of endotoxin (de Rooij et al. 2017; Mueller-Anneling et al. 2004; Pavilonis et al. 2013a; Schulze et al. 2006; Tager et al. 2010). Study limitations hampered identification of livestock-related determinants. Until now, well-designed large-scale measurement studies assessing spatial variation of endotoxin have been lacking, largely due to the difficulties and costs associated with endotoxin measurements. Measurement of the mass of particulate matter (PM) $\leq 10\,\mu m$ ($PM_{10}$) is less challenging and provides opportunities like real-time assessment. It is yet unknown whether $PM_{10}$ mass measurements represent spatial variation of livestock endotoxin emissions well in a livestock-dense area that is without other significant PM sources.

In-depth exposure assessment studies based on airborne measurements focusing on urban air pollution related to long-term health effects have been performed for many years. A widely used technique in this field is land-use regression (LUR) modeling (Beelen et al. 2013; de Hoogh et al. 2013; Eeftens et al. 2012a; Jerrett et al. 2005; Ryan and LeMasters 2007), a modeling approach that uses geospatial predictor variables to explain spatial contrasts in measured airborne concentrations. Recently this technique has also proven useful in an urban environment for bioaerosols, namely allergenic pollen (Hjort et al. 2016).

We repeatedly measured concentrations of endotoxin in the $PM_{10}$ fraction in a livestock-dense area in Netherlands at 61 sites representing a variety of livestock-related characteristics. Our first aim was to assess spatial variation of endotoxin and $PM_{10}$ in a livestock-dense area. The second aim was to evaluate how well exposure proxies used in epidemiological studies relate to measured concentrations. Our third aim was to assess whether the LUR approach can be used for modeling measured endotoxin concentrations in relation to livestock-related characteristics of the surroundings in a rural area.

## Methods

### Strategy

The study area was situated in the southeast of Netherlands. A large epidemiological study was started in 2012 in that region aimed at assessing the health of residents living in livestock-dense areas (Borlée et al. 2015, 2017). Measurement sites were selected to represent the areas of residence of the health study participants (Figure 1). From May 2014 to December 2015, ambient $PM_{10}$ was collected repeatedly at 61 sites in residential gardens. The measurement strategy was based on previous LUR campaigns focusing on urban air pollution using a spatially dense network of measured concentrations (Hoek et al. 2008). Measurements were performed simultaneously at 10 sites per run. The aim was to collect four 2-wk averaged samples per site in different seasons. A reference site in the study area was set up where measurements were taken continuously to account for temporal variability. This enabled calculation of annual average concentrations for all sites, as previously reported (de Hoogh et al. 2013; Eeftens et al. 2012a). Geographical information system (GIS) software (version 10.2.2; ArcGIS) was used to compute general and detailed livestock characteristics of the surroundings. Measurement sites were geocoded and plotted in ArcGIS in combination with the geocoded livestock data. Coordinates of the locations of all livestock farms and the number and species of licensed animals per farm in the year 2015 were used. These were obtained from the provincial database of mandatory environmental licenses for livestock keeping, as provided by the provinces of Noord-Brabant (http://bvb.brabant.nl) and Limburg (http://limburg.vaa.com/webbvb). LUR models for annual average concentrations of endotoxin and $PM_{10}$ were developed based on livestock-related GIS predictors by using a supervised step-wise selection procedure for LUR modeling described previously (Eeftens et al. 2012a).

### Study Area

The study area ($3,000\,km^2$ in size) comprised regions of the Netherlands with the highest livestock density. The study area was situated in the provinces of Noord-Brabant and Limburg. The number of farms in Noord-Brabant and Limburg is 13,670 and 3,490, respectively, on a surface of $5,081\,km^2$ and $2,209\,km^2$, respectively, (provincial databases, http://bvb.brabant.nl and http://limburg.vaa.com/webbvb). Farms are not equally distributed; instead they are predominantly concentrated in the region where the two provinces border (Figure 1). In Netherlands, livestock is commonly kept in enclosed animal houses (average size of premises 12.8 ha in 2015), apart from some dairy cows, sheep, and horses that also are kept on pastures during parts of the year (average size of premises 31.9 ha in 2015) (Centraal Bureau voor de Statistiek 2017a). Sizes of farm premises are highly variable, as are the number and size of animal houses per farm, which also depend on the animal species kept. Most farms are specialized: Only one animal species is kept, focusing on a specific product (e.g., broiler farms, laying hen farms). The animal numbers kept on a farm vary greatly; on average, the numbers kept on pig farms and chicken farms were 2,600 and 52,000 respectively, in 2015 (Centraal Bureau voor de Statistiek 2017b).

### Site Selection

Measurement sites were selected to cover a wide spatial contrast in livestock farm density and animal species kept (see the "GIS Predictors" section for details) in the rural areas of residence of the health study participants. Four categories were made based on the distance from the site to the nearest farm: <250 m, 250–500 m, 500–1,000 m, and >1,000 m. We chose to overrepresent sites close to livestock farms to ensure coverage of the expected highest local-scale air pollution variation in the vicinity of farms. Sites were distributed as follows: 40% were <250 m, 35% 250–500 m, 18% 500–1,000 m, and 7% >1,000 m. Furthermore, attention was paid to ensure that farm numbers of the main animal species (cattle, poultry, and swine) had a distribution without major outliers over the different sites. The reference site was selected according to criteria previously described for urban LUR campaigns: It was located in the same study area and not directly influenced by local sources (Hoek et al. 2008). The distance from the reference site to the nearest farm was 1,200 m. At this location, two $PM_{10}$ samplers were placed side-by-side for simultaneous measurements. These parallel measurements were used to gain insight into the correlation between parallel samples and to increase the precision of temporal adjustments by taking averages.

Residents were visited by fieldworkers, informed about the study, and invited to participate. To limit the influence of non-livestock local $PM_{10}$ sources, residences within 500 m of a highway, a railway, or an industrial area and 50 m of a busy provincial/interurban road were not eligible for inclusion. Inspection of the premises of those interested in participation was carried out to ensure the location was suitable for the sampling equipment. The equipment was placed as far away as possible from the nearest road. According to the Central Committee on Research involving Human Subjects (CCMO), this type of study does not require approval from an ethics committee in the Netherlands.
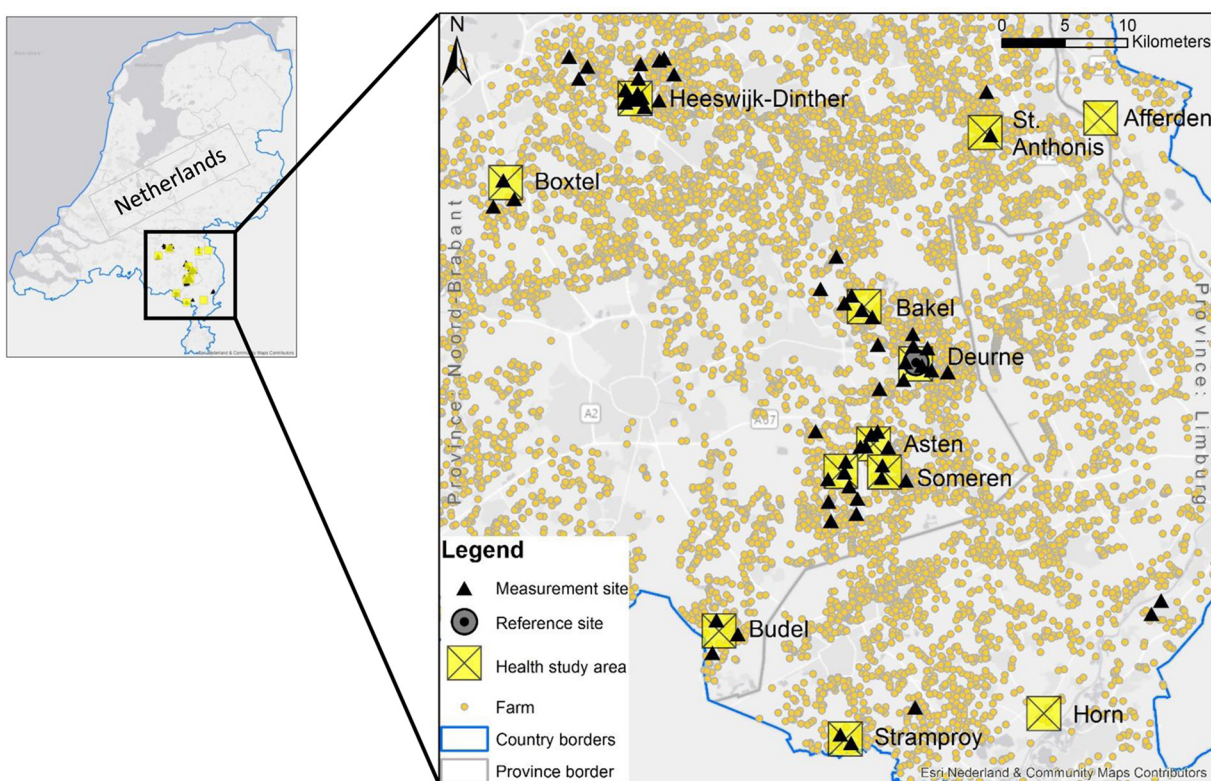
**Figure 1.** Geographical distribution of the 61 measurement sites in a livestock-dense area in Netherlands covering the geographical spread of the health study areas. Note: Health study area: area where participants of the VGO study (Livestock Farming and Neighboring Residents' Health Study) lived, see Borlée et al. 2017 for maps of the study area depicting exact localization of residences. The figure was generated using ArcGIS (version 10.2.2; Esri) and the gray background was sourced from Esri Nederland & Community Maps Contributors.

### Measurements

Harvard impactors (Air Diagnostics and Engineering Inc., Naples, ME, USA) were used to sample $PM_{10}$ on Teflon filters (Teflo W/ ring 37 mm with 2-μm pore size; SKC, PA, USA). Samples were taken at a height of 1.6 m, the average breathing height of humans. A self-designed pump unit was used, and the flow was maintained at 10 L/min by means of critical orifices, as described previously (Eeftens et al. 2012b; see the European Study of Cohorts for Air Pollution Effects (ESCAPE) project website (http://www.escapeproject.eu/manuals/) on the standard operating procedure for PM). The air flow was measured before and after sampling using a calibrated rotameter. Pumps were installed to sample 15 min of each hour during the 14-d period to avoid filter overloading. Total run time of the pump was recorded by elapsed time counters. Total sampling volume was calculated based on the total sampling time and the average flow.

The aim was to sample each site successfully for four times distributed over the four seasons. If a measurement at a site failed, it was repeated in the next round. Failures could be due to loss of power, a blocked pump, a damaged filter, or extreme weather conditions. In addition, we excluded samples with a start or end flow of $<8.5$ L/min and samples that did not fulfill sampling duration criteria (within 33% margin of 84 h). Each measurement period, a field blank control was placed at a different site. This field blank filter underwent the same procedure except that no air was drawn through the sampling device. All samples were stored within 72 h after collection at $-20°C$.

### Laboratory Analyses

$PM_{10}$ mass was determined by gravimetric analysis using a microbalance (1-μg precision), following the protocol described

by Eeftens et al. (2012b) [see ESCAPE project website (http://www.escapeproject.eu/manuals/) on the standard operating procedure for weighing]. Briefly, filters were conditioned for 24 h prior to pre- and post-weighing in a weighing room with controlled temperature ($21 \pm 0.5°C$) and relative humidity ($35 \pm 5\%$). The limit of detection (LOD) was defined as three times the standard deviation of the field blanks. Subsequently, samples were processed for endotoxin extraction and analysis as described by Spaan et al. (2008a). In short, filters were transferred to 50-mL tubes (Greiner Bio-one) and 5 mL of pyrogen-free water (Aqua B. Braun) supplemented with 0.05% Tween 20 (Calbiochem, USA). After shaking for 1 h in an end-over-end roller, tubes were centrifuged at $1,000 \times g$ for 15 min and 1 mL of supernatant was stored at $-20°C$. Endotoxin was analyzed by means of a limulus amebocyte lysate (LAL) assay, and Tween 20 was not used in the assay solution as per recommendations by Spaan et al. (2008a). Briefly, 100 μL of 25-times diluted sample was used in the quantitative kinetic chromogenic LAL assay (Lonza, Walkersville, MD, USA; LAL-lysate lot number PL067KXRWM). For the endotoxin assay, the LOD was defined as the smallest concentration that produced a signal in the calibration curve. Results are expressed as endotoxin units (EU) per cubic meter of sampled air.

### Adjustment for Temporal Variation

The reference site was used for adjustment of temporal variation over the measurement period using the difference method as described earlier (de Hoogh et al. 2013; Eeftens et al. 2012a). Briefly, we calculated the difference in $PM_{10}$ and endotoxin concentrations measured at the reference site during each 2-wk measurement period from average $PM_{10}$ and endotoxin concentrations at the reference site over the entire study period (May

2014–December 2015). We then subtracted this difference from the $PM_{10}$ and endotoxin concentrations measured at each individual (nonreference) site during the same 2-wk period. Annual average $PM_{10}$ and endotoxin concentrations at each site were computed as the arithmetic mean of the reference site-adjusted concentrations for each site. Imputations were performed for $PM_{10}$ mass concentration at the reference site for the first measurement period—thus at the start of the measurement campaign in May 2014—because the installations at the reference site functioned well during the whole measurement period except for the first measurement period. $PM_{10}$ mass concentrations measured at the reference site were highly correlated (Pearson correlation 0.92, $p < 0.001$) with $PM_{10}$ mass concentrations measured at a national monitoring station situated at a 10-km distance (see Figure S1). Based on $PM_{10}$ levels measured at the national monitoring station during the first measurement period, the $PM_{10}$ level was imputed for the reference site for that period. No imputation for endotoxin was performed because the variation of endotoxin concentrations at the reference site per period could not be predicted from air monitoring data given that endotoxin is not routinely monitored.

### GIS Predictors

The number of animals and the number of farms and specific farm types (cattle, pigs, poultry, goats, sheep, horses, fur animals) in buffer zones of respectively 250, 500, 1,000, 3,000 m around the sites were computed (Table 1; see also Table S1). Buffer size was restricted to 3,000 m because of proximity to country borders; no livestock data was available for regions outside Netherlands. We calculated the distance from each measurement site to the nearest farm of each type, distance was taken into account linear and inversed because the shape of the potential association was unknown. We used circular buffers with the number of farms or animals summed up in that buffer. To better account for the dilution effect with increasing distance from the source, we weighted the number of animals per species and the number of farms per farm type within a 1,000-m or 3,000-m buffer by the distance of each farm from the measurement site.

### Statistical Analyses

Linear regression analyses were performed on annual average concentrations of $PM_{10}$ and endotoxin with multiple GIS-derived livestock-related variables as input. Univariable analyses were performed to compare simple livestock exposure proxies to measured endotoxin concentrations. To avoid models heavily affected by a few observations, and to ensure enough power to analyze the GIS-derived predictors, predictors with a value of zero for

>40 of the 61 sites (two-thirds of total number of sites) were excluded from multivariable regression analyses. Each GIS variable was truncated to its 95th percentile value. This was done because various GIS predictors showed a right-skewed distribution among monitoring sites and because sensitivity analyses using nontruncated predictors showed multiple modeling results to be heavily affected by a few sites (data not shown).

Concentrations of $PM_{10}$ and endotoxin approximated a normal distribution. GIS variables were selected using a forward supervised stepwise selection procedure (Eeftens et al. 2012a). First, the direction of the effect for the variables was determined (Table 1). Then, with each step, the variable in the predefined direction with the highest gain in adjusted $R^2$ was entered in the model on condition that addition of the variable did not change the direction of already-included variables. Model building was terminated when inclusion of a variable did not improve the adjusted $R^2$. Model assumptions were checked for the resulting models including distribution of residuals and spatial autocorrelation through Moran's $I$ statistic. Additionally, variables were removed if their $p$-value was >0.10 and/or there was evidence of collinearity (variance inflation factor >3). Cook's Distance was checked for all model observations and examined further if higher than 1. Because of the widely different scales of the different livestock-related variables assessed, each variable was scaled from the 10th percentile to 90th percentile range, making direct comparisons of the magnitude of effects of GIS variables possible.

The LUR modeling procedure was performed with different levels of detail of livestock-related variables. First, an LUR model was developed based on general livestock-related characteristics (distance to nearest farm, number of farms within buffer zones, and weighted distance number of all farms) because these have been used frequently as exposure proxies in epidemiological studies. Then, LUR modeling (procedure 2) was based on a larger predictor set including more refined GIS predictors: animal species–specific variables (farms and animal numbers of cattle, poultry, swine, goats, sheep, horses, and fur animals) that were presented in addition to the general livestock-related characteristics. Finally, LUR modeling (procedure 3) was carried out based on all available GIS predictors, thus including farm type–specific information for cattle, poultry and swine farms (dairy cattle, beef cattle; broilers, laying hens, other; piglets, fattening pigs, sows).

Model development using a high number of predictor variables has the risk of overfitting (Babyak 2004); therefore, several precautionary measures were applied before and during modeling (i.e., high number of sites, predetermined direction of effect, limit of nonzero values for predictors, tiered model procedure). In addition, after models were developed, validation methods were applied in order to gain insight into possible overfitting. An internal validation

**Table 1.** Overview of the different types of livestock-related predictor variables, different buffer sizes, and determined direction of effect.

| Predictor | Categories | Unit | Buffer sizes (m) | Direction of effect |
|---|---|---|---|---|
| Number of animals in a buffer | Cow, pig, poultry, goat, sheep, horse, fur animals | $n$ | 250, 500, 1,000, 3,000 | + |
| Number of farms in a buffer | All, cow (dairy, meat), pig (piglets, fattening pigs, sows), poultry (laying hens, broiler, other), goat, sheep, horse, fur animals | $n$ | 250, 500, 1,000, 3,000 | + |
| Distance to nearest farm (linear and inversed) | Any, cow, pig, poultry, goat, sheep, horse, fur animals | $-1^*$m and m$^{-1}$ | NA | + |
| Number of animals in a buffer weighted for distance to site | Cow (dairy, meat), pig (piglets, fattening pigs, sows), poultry (laying hens, broiler, other), goat, sheep, horse, fur animals | $\Sigma\,(n/m)$ | 1,000, 3,000 | + |
| Number of farms in a buffer weighted for distance to site | All, cow (dairy, meat), pig (piglets, fattening pigs, sows), poultry (laying hens, broiler, other), goat, sheep, horse, fur animals | $\Sigma\,(n/m)$ | 1,000, 3,000 | + |

Note: Some variables were computed to be zero for all sites, meaning that this type of farm/animal was not present in the surroundings (defined by the buffer size taken into account) around the measurement sites. NA, not available.

method, leave one out cross validation (LOOCV), was applied to evaluate overall model performance as in previous LUR studies (Eeftens et al. 2012a). This method entails sequentially leaving out each site from the model while the included set of predictors are left unchanged. Model robustness was further assessed by hold-out validation (HV) because this method has been shown to be a more stringent validation test (Basagaña et al. 2012; Wang et al. 2012). Ten-fold HV was performed using random selections of 90% of the measurement sites. Additionally, we performed a sensitivity analysis for LUR modeling procedure 3 by applying less restrictive requirements with respect to the numbers of sites having a nonzero value: 15% of the sites (instead of one-third of the sites) having a nonzero value for each predictor. This way the effect of taking into account smaller (250 m, especially) buffer sizes for the various animal species was investigated.

## Results

### Descriptives of Measurements

In total, 236 successful measurements were performed distributed over 61 sites (number of sites measured for either three, four, or five times was 10, 49, and 2, respectively). The installations at the reference site functioned well during the whole measurement period except for the first measurement period. Imputation was performed for $PM_{10}$ but not for endotoxin because the variation of endotoxin concentrations at the reference site per period could not be predicted from air monitoring data given that endotoxin is not routinely monitored. For annual average computations of endotoxin concentrations, measurements performed at sites in the first period were not taken into account (computations based on 229 measurements; number of sites measured for either three, four, or five times was 16, 44, and 1, respectively). For $PM_{10}$, all 236 successfully performed measurements could be used for annual average computations. For both endotoxin and $PM_{10}$ 23 of 24 sampling blanks were below the LOD; whereas the remaining blank was just above the LOD for endotoxin as well as the LOD for $PM_{10}$. Because concentrations of all sampled filters were well above the LOD, no adjustment for blanks was applied.

Two-week average endotoxin as well as $PM_{10}$ mass concentrations showed clear variation over time (Figure 2). Endotoxin and $PM_{10}$ temporal variation patterns differed illustrated by the low correlation between endotoxin and $PM_{10}$ mass concentrations at the reference site (Pearson correlation 0.19, $p = 0.37$). The coefficient of variation (CV) between side-by-side collected parallel samples at the reference site was higher for endotoxin (mean CV 17.7%, Pearson correlation 0.68, $p < 0.001$) than for $PM_{10}$ (mean CV 1.9%; Pearson correlation 0.99, $p < 0.001$). Samples analyzed repeatedly for endotoxin in the laboratory on different days showed high correlation and limited variability between repeats (30 samples; mean CV 13.7%; Pearson correlation 0.95, $p < 0.001$). This documents the higher inherent variability for endotoxin compared with $PM_{10}$ due to both sampling and analytical variability. Endotoxin and $PM_{10}$ mass concentrations measured at the reference site during each 14-d period were highly correlated with endotoxin and $PM_{10}$ mass concentrations averaged over the sites measured during the same time period (Pearson correlation 0.75, $p < 0.001$; 0.98, $p < 0.001$, respectively). Endotoxin concentrations were generally lower at the reference site compared with the average concentration of sites measured during the same measurement period (Figure 2).

### Spatial Variation

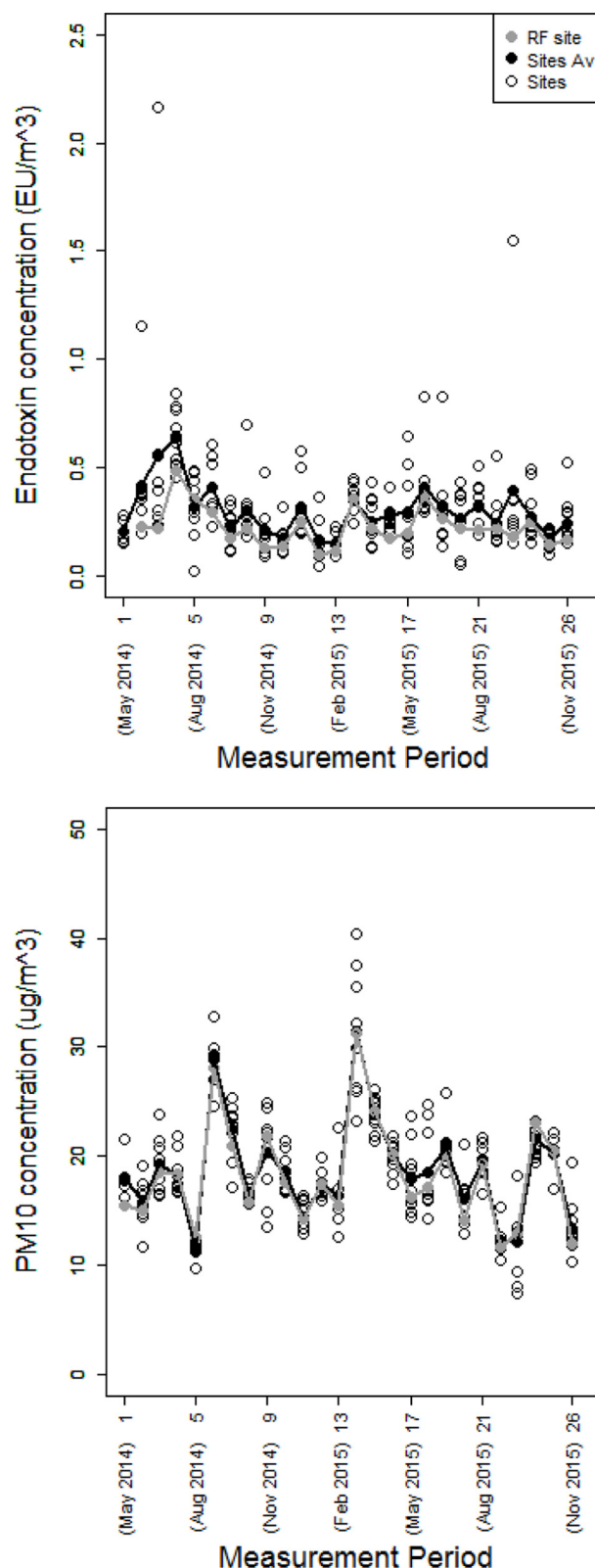Annual average endotoxin concentrations ranged from 0.13 to 0.85 EU/m³ (6.5-fold difference) between sites, $PM_{10}$ mass



**Figure 2.** Overview of endotoxin (EU/m³) and $PM_{10}$ (μg/m³) concentrations measured during multiple 2-wk periods at 61 measurement sites, and continuously at a reference (background) site, in a livestock-dense area. Note: Sites avg, average concentration over all measurement sites during the measurement period; RF site, reference site; sites, measurement sites.
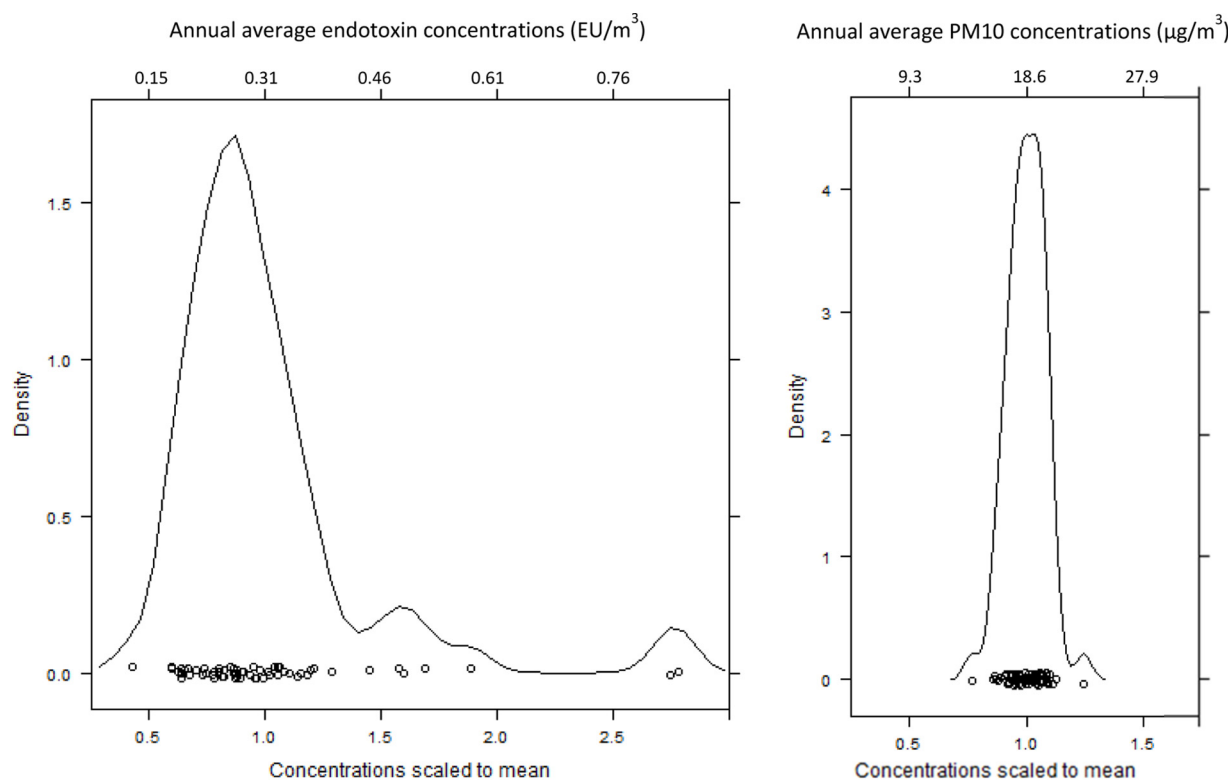
**Figure 3.** Density plots depict variation in annual average concentrations of endotoxin and $PM_{10}$ between 61 sites measured in a livestock-dense area.

concentrations from 14.4 to 23.1 µg/m$^3$ (1.6-fold difference). More variation in annual average concentrations between sites was observed for endotoxin compared with $PM_{10}$ (between sites CV 43%, 8%, respectively), Figure 3. Correlation between annual average concentrations of endotoxin and $PM_{10}$ was moderate (see Figure S2). Table 2 shows descriptive statistics of measured concentrations at sites categorized by distance to nearest livestock farm. Higher mean values for endotoxin were observed among sites within the closer distance categories. Endotoxin concentrations varied considerable among sites within the same distance category (e.g., from 0.18 to 0.85 EU/m$^3$ for the 24 sites <250 m from a farm) (Table 2). Mean $PM_{10}$ mass concentrations did not differ between the distance categories.

### Exposure Proxies

Univariable linear regression models showed a high number of livestock-related GIS variables potentially associated with endotoxin concentrations (see Table S2). All of the general livestock characteristics (distance to nearest farm, number of farms within buffer zones, and weighted distance number of all farms) were associated with increased endotoxin concentrations, but none explained >26% of the spatial variation in endotoxin concentrations (Table S2). Of the species-specific variables, pig- and poultry-related variables were the most strongly related to

endotoxin concentrations. Of all univariable analyses, the parameter distance-weighted number of sows in a 1,000-m buffer explained the largest amount of the variance in endotoxin concentrations ($R^2 = 0.35$). This variable was highly correlated with distance-weighted total number of pigs within a 1,000-m buffer (Pearson correlations 0.77) but was not strongly correlated with distance-weighted numbers of other animal species (Pearson correlations −0.03 to 0.34) (see Figure S3 and Figures S4–S6 for correlations among distance-weighted numbers of animals within 1,000 m/3,000 m and numbers of farms within 1,000 m and 3,000 m, respectively).

### LUR Models

The three LUR models for endotoxin resulting from modeling procedures including successively more detailed livestock-related predictors are described in Table 3. Model residuals met the criteria of normality, homoscedasticity, and spatial independency (Moran's $I$ was not significant). The endotoxin LUR model based on general livestock characteristics only (Model 1), included two predictors (the number of farms in a 250-m buffer and the distance-weighted number of farms in a buffer of 1,000 m) that explained 30% of spatial variation (LOOCV $R^2 = 0.19$, 10-fold HV $R^2 = 0.10$). Model 2, based on the modeling procedure incorporating also animal species–specific predictors, included three

**Table 2.** Overview of measured annual average concentrations of endotoxin and $PM_{10}$ at the sites distributed over four distance categories related to distance to closest farm.

| Distance to nearest livestock farm | Number of sites | Endotoxin (units/m$^3$) | | | | $PM_{10}$ (µg/m$^3$) | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | AM | SD | Minimum | Maximum | AM | SD | Minimum | Maximum |
| <250 m | 24 | 0.36 | 0.17 | 0.18 | 0.85 | 19.1 | 1.5 | 16.7 | 23.1 |
| 250–500 m | 21 | 0.29 | 0.11 | 0.2 | 0.58 | 18.4 | 1.5 | 14.4 | 20.3 |
| 500–1,000 m | 11 | 0.25 | 0.04 | 0.19 | 0.31 | 18 | 1.4 | 16.1 | 19.9 |
| >1,000 m | 5 | 0.24 | 0.06 | 0.13 | 0.29 | 18.9 | 1.1 | 17.7 | 20.5 |

Note: AM, arithmetic mean; SD, standard deviation.

**Table 3.** Endotoxin LUR models resulting from three modeling procedures: *a*) considering general livestock variables only, *b*) considering general plus animal species–specific variables, *c*) considering general plus animal species–specific plus farm type–specific variables.

| Model | Variables (scaled to 10–90th percentile range) | 10th Percentile | 90th Percentile | Estimate (90% CI) |
|---|---|---|---|---|
| Model 1 | Intercept | – | – | 0.177 (0.094, 0.258) |
| Model $R^2$: 0.30 | Number of livestock farms in a 250-m buffer | 0 | 2 | 0.091 (0.033, 0.148) |
| Model adj. $R^2$: 0.27 | Number of livestock farms weighted to distance in | 0.003 | 0.041 | 0.091 (0.007, 0.175) |
| LOOCV $R^2$: 0.19 | a 3,000-m buffer [$\Sigma$ ($m^{-1}$)] | | | |
| HV $R^2$: 0.10 | | | | |
| Model 2 | Intercept | – | – | 0.082 (0.003, 0.160) |
| Model $R^2$: 0.48 | Inversed distance to nearest pig farm ($m^{-1}$) | 0.001 | 0.006 | 0.169 (0.111, 0.227) |
| Model adj. $R^2$: 0.45 | Number of poultry animals weighted to distance in | 42.573 | 1014.177 | 0.112 (0.050, 0.173) |
| LOOCV $R^2$: 0.37 | a 3,000-m buffer [$\Sigma$ ($n$/distance in m)] | | | |
| HV $R^2$: 0.22 | Number of horse farms in a 3,000 m buffer | 8 | 18 | 0.065 (0.006, 0.123) |
| Model 3 | Intercept | – | – | 0.122 (0.057, 0.186) |
| Model $R^2$: 0.64 | Number of sows weighted to distance in a 1,000-m | 0 | 5.803 | 0.148 (0.107, 0.189) |
| Model adj. $R^2$: 0.61 | buffer [$\Sigma$ ($n$/distance in m)] | | | |
| LOOCV $R^2$: 0.51 | Number of laying hens weighted to distance in a 3,000-m | 3.465 | 376.545 | 0.114 (0.062, 0.167) |
| HV $R^2$: 0.32 | buffer [$\Sigma$ ($n$/distance in m)] | | | |
| | Number of poultry animals in 500-m buffer | 0 | 77,136 | 0.086 (0.040, 0.131) |
| | Number of horse farms in a 3,000-m buffer | 8 | 18 | 0.063 (0.014, 0.111) |

Note: Endotoxin concentrations were annual average concentrations ($EU/m^3$) measured at 61 sites in a livestock-dense area. Model 1 resulted from the modeling procedure taking into account only general livestock variables; Model 2 resulted from the modeling procedure taking into account general plus animal species–specific variables; Model 3 resulted from the modeling procedure taking into account general plus animal species–specific plus farm type–specific variables. See Table S2 for a complete list of variables included in each group. Predictor variables were truncated to the 95th percentile and then scaled to the 10–90th percentile range, thus predictor values were divided by the 10–90th percentile range of that predictor. –, no data; CI, confidence interval; HV, hold-out validation; LOOCV, leave one out cross validation; Model adj. $R^2$, model adjusted $R^2$.

predictors (related to pigs, poultry, and horses, respectively) and had a considerably higher amount of explained spatial variation (48%) (LOOCV $R^2 = 0.37$, 10-fold HV $R^2 = 0.22$). Model 3, based on the modeling procedure that additionally incorporated farm type–specific predictors, included four predictors and explained even more spatial variation compared with Model 2, with an $R^2$ of 64% (LOOCV $R^2 = 0.51$, 10-fold HV $R^2 = 0.32$). Different types of predictors were considered in the models, including animal densities, farm densities, and distance-weighted characteristics. The same animal species (pigs, poultry, and horses) were included in Model 3 as in Model 2.

As for endotoxin LUR models, $PM_{10}$ LUR model performance improved as more specific livestock-related predictors were taken into account. $PM_{10}$ LUR models included different and fewer predictors (1 for Models 1 and 2, 2 for Model 3) and the spatial variance explained was lower than for the endotoxin LUR models ($R^2$ values of 0.08, 0.14, and 0.19 for Models 1, 2, and 3, respectively) (see Table S3).

In a sensitivity analysis, we repeated endotoxin and $PM_{10}$ modeling considering general, species-specific, and farm type–specific predictors (modeling procedure 3) but allowed predictors with only 15% of sites with nonzero values (vs. one-third in the primary analyses) (Table S4). The resulting endotoxin LUR model included one new predictor (the number of pigs in a 250-m buffer), in addition to the four predictors in the primary model, and had a slightly higher $R^2$ (0.66 vs. 0.64). The $PM_{10}$ LUR model included three predictors, none of which were in the primary model, and had a higher $R^2$ (0.31 vs. 0.19).

## Discussion

This study, with 61 measurement sites distributed over a livestock-dense area, shows clear spatial variation in annual average endotoxin concentrations associated with livestock density. Thus, residential exposures to endotoxin are variable within one region, providing great opportunities for LUR modeling. To our knowledge, LUR modeling as described in this study has not been applied for endotoxin before. The LUR models developed explain spatial variation of endotoxin concentrations to a large extent. General livestock characteristics, such as livestock density or distance to the nearest farm, which

have been used as exposure proxies in previous epidemiological studies (Borlée et al. 2015; Elliott et al. 2004; Mirabelli et al. 2006; Pavilonis et al. 2013b; Radon et al. 2007; Sigurdarson and Kline 2006; Smit et al. 2014; Wing and Wolf 2000), were positively and statistically significantly associated with endotoxin concentrations, although spatial variation explained was limited. Models using a combination of animal species-specific and farm type–specific predictors performed notably better. Mean endotoxin concentrations measured at sites <250 m from a livestock farm were 50% higher than mean concentrations measured at sites >1,000 m from a farm (0.36 vs. 0.24 $EU/m^3$, Table 2), indicating a sizable contrast in exposure. The large variability in endotoxin concentrations within distance categories suggests the importance of farm characteristics (size, animal species). Annual average concentrations of endotoxin were not represented well by $PM_{10}$. Annual average $PM_{10}$ mass concentrations showed more limited spatial variation and a different spatial pattern. $PM_{10}$ is, compared with endotoxin, likely a less specific pollutant for livestock farming and $PM_{10}$ mass concentrations in the studied area are characterized by a high background. The low spatial variability of $PM_{10}$ is comparable to studies focusing on urban settings, where the spatial variation of $PM_{10}$ and $PM_{2.5}$ (PM ≤2.5 μm) is also lower than more source-specific components of PM such as elemental carbon (Eeftens et al. 2012a; Jedynska et al. 2014).

### Concentrations Measured

Until now, only measurement studies including a limited number of sites have been performed to characterize airborne endotoxin in the rural environment (de Rooij et al. 2017; Heederik and Ijzermans 2011; Mueller-Anneling et al. 2004; Pavilonis et al. 2013a; Schulze et al. 2006; Tager et al. 2010). Measured concentrations were generally comparable (within the same order of magnitude) for the different studies; however, a detailed comparison is not possible because of differences in methodology, including sampling of different particle size fractions (majority $PM_{10}$, one study including coarse fraction (Tager et al. 2010), and one study including inhalable fraction (Schulze et al. 2006) averaging times and sampling strategies. Two small studies (one including five sites, the other eight sites) using similar methods as in

the current study reported ranges of 0.21–0.31 EU/m$^3$ and 0.46–0.66 EU/m$^3$, respectively (de Rooij et al. 2017; Heederik and Ijzermans 2011). Spatial variation measured between rural sites in these studies was less compared with the current study, which included a high number of sites (61) and specifically focused on sampling of a variety of sites in a livestock-dense area.

Agreement and correlation of parallel, side-by-side, measured endotoxin concentrations at the reference site was considerable but not close to 1 as for parallel-measured PM$_{10}$ levels. This is similar to what we observed previously (de Rooij et al. 2017). Analytical variability only partly explains the differences observed for parallel collected samples. The residual variability could be caused by high very local-scale (within meters) variation for endotoxin. This is likely related to the profound influence of local sources with variable emission strength and to the endotoxins' origin: bacteria, which can coagulate, grow, and amplify (de Rooij et al. 2017). Moderate side-by-side correlation has implications for assessment of the model $R^2$; model $R^2$ should in that case not be set against 100%. That is, the moderate side-by-side correlation measured for endotoxin will reduce the percent explained variance of models based on endotoxin measurements.

### Exposure Proxies in Epidemiology

Health surveys that explored livestock-associated health effects have used simple and general livestock characteristics as exposure proxies, including farm density in a buffer and various distance measures (Borlée et al. 2015; Elliott et al. 2004; Pavilonis et al. 2013b; Radon et al. 2007; Smit et al. 2014; Wing and Wolf 2000). All general livestock characteristics explored in this study were significantly associated with measured endotoxin concentrations. In particular, spatial variation explained by the number of farms in buffers was encouraging ($R^2 \leq 0.26$, given the relative simplicity and crudeness of these proxies. This provides evidence of the usefulness of general livestock characteristics used as exposure proxies. However, explained spatial variation was modest, particularly for the simple distance metrics, leaving room for improvement of endotoxin exposure estimation. The restrictions of simple distance metrics were also demonstrated in the study by Cantuaria et al. (2016), in which different approaches for assessing ammonia exposure in rural areas were compared. Models that improve exposure estimation should be preferred over the use of proxies as exposure misclassification can give rise to biased results and/or loss of precision (Armstrong 1998).

### LUR Models

To our knowledge, this study is the first to apply an LUR approach for endotoxin in a rural area, showing that spatial variation of annual average endotoxin concentrations can be explained well by (refined) livestock information. The explained spatial variation of the endotoxin LUR model based on refined livestock information was high, especially when taking into account the moderate correlation of parallel samples. The better performance of models with refined livestock predictors is in line with results from studies measuring on farms, in which concentrations of endotoxin differ between animal species kept but also between specific farm type (i.e., broilers vs. laying hens) and farm size (number of animals) (Ogink et al. 1997; Seedorf et al. 1998).

All endotoxin models contained distance and density of animal/farm metrics underlining the importance of the number in combination with proximity of sources. An array of buffer sizes was taken into account in the modeling to ensure capturing relevant distances. Predictors eventually included in the models represented small buffer as well as larger buffer sizes, reflecting the

influence of farms at close distances and the collective impact of farms in the area. Predictors related to the animal species poultry and pigs prevailed. Considering that LUR models are not source attribution models, these results are in line with previous studies suggesting substantial endotoxins emissions of especially poultry farms but also pig farms (Seedorf et al. 1998). Animal species represented in the LUR model, which included all available livestock predictors (including farm type–specific information), were highly comparable to simpler models indicating good modeling consistency. Effect estimates for the in Model 3 included variables (scaled to 10–90th percentile) were comparable to each other, none had a strikingly outlying value. Contributions to ambient concentrations of a specific animal type depend on individual source strength of a farm and the number of those farms present. In our study area, the number of poultry farms was considerably lower compared with pig farms and cattle farms; however, poultry farms are known to have high emissions (Seedorf et al. 1998). The combination of variation in numbers present of specific farm types and variation in emission levels of these farm types might explain the relatively similar effect estimates.

In contrast to the endotoxin LUR models, the LUR models for PM$_{10}$ were less successful in explaining spatial variation of annual average concentrations in relation to livestock information. PM$_{10}$ mass concentrations were more poorly associated with livestock presence, few associated livestock-related characteristics were found, especially in the smaller buffer sizes. Spatial variation could not be explained by an LUR model developed earlier for urban air pollution exposure as part of the ESCAPE project (Eeftens et al. 2012a) for Netherlands/Belgium ($R^2$ 0.003 for this model used on our 61 sites). This is probably explained by the rural character of the study area and our sampling strategy in which traffic-affected sites were avoided to focus on PM$_{10}$ variability related to livestock emissions.

### Model Evaluation

Model evaluation is important as LUR modeling has the risk of overfitting with potentially a large number of predictor variables taken into account to explain concentrations at relatively few sites. To limit the risk of overfitting several precautionary measures were applied before and during modeling including inclusion of a large number of sampling sites. The number of 61 sites included in the current study is generally considered sufficient for LUR modeling of urban predominantly traffic-related air pollution studies (Hoek et al. 2008; Ryan and LeMasters 2007). The number of predictor variables offered was highest in the most refined model. The difference in $R^2$ between the three LUR models, however, does not simply reflect model overfitting by offering more predictor variables in the modeling procedure because the same pattern in performance was found for validation statistics.

The gap between endotoxin LUR model $R^2$ and $R^2$ of LOOCV and especially HV was larger than expected based upon the number of sites included. Previous studies evaluating the effect of number of sites on nitrogen dioxide (NO$_2$) LUR model validation reported smaller differences (Basagaña et al. 2012; Wang et al. 2012). This might be explained by more substantial variation at source level for endotoxin. Seedorf et al. (1998) showed not only that indoor endotoxin levels differ between farms of the same type within the same country, but also that differences within the same farm at different time points can be substantial. Other studies measuring endotoxin concentrations on farms and/or at farm premises also detected considerable differences in endotoxin concentrations over time, suggesting varying endotoxin emissions (Ogink et al. 1997; Thorne et al. 2009). Studies on PM$_{10}$ emissions found similar results and differences

between and within farms were suggested to be related to, among other factors, production cycle and farm management (Winkel et al. 2015). Livestock data incorporated involved farm location, farm type, and permitted animal numbers. Information related to specific individual farm characteristics was not available. The effect of the higher source level variability for endotoxin might increase even more as a result of geographic uncertainty on farm emission points. Consequently LUR modeling of endotoxin concentrations on the basis of livestock data can be regarded as more challenging than modeling of air pollution in relation to urban predominantly traffic-related sources.

### Health Implications

The long-term average concentrations were higher at sites with more livestock production in the surroundings, but markedly below concentrations known to cause endotoxin-related health effects. Occupational studies do not report health effects for concentrations observed in this study (Basinas et al. 2015; Liebers et al. 2006; Samadi et al. 2013). However, direct inferences cannot be made because results from those studies specifically apply for a short duration of exposure (around 4–8 h) in a healthy worker population. Moreover, different size fractions (mainly inhalable or respirable dust) were sampled in those studies, and it is unknown how health effects can be allocated to specific particle size fractions. Identified elevated long-term average concentrations in our study are likely associated with a higher frequency and level of peak exposures because emissions from farms are not constant and distribution depends on, among other factors, atmospheric conditions including wind direction, speed, atmospheric stability, and mixing layer height. Results of studies measuring short-term (several hours) at farm premises indicate the occurrence of such peaks: concentrations of $30\,EU/m^3$ and far above were measured (Jonges et al. 2015; Thorne et al. 2009). More knowledge on exposure levels over different averaging times and different particle size fractions in relation to health effects in the general population is urgently required.

### Study Limitations

Endotoxin concentrations were measured in the $PM_{10}$ fraction, no information was obtained on endotoxin concentrations in other size fractions. Studies measuring ambient endotoxin in $PM_{2.5}$ and $PM_{10}$ showed higher concentrations in the $PM_{10}$ fraction (Chen and Hildemann 2009; Morgenstern et al. 2006; Pavilonis et al. 2013a; Soukup and Becker 2001), suggesting underestimation of concentrations when assessing particle size fraction smaller than $PM_{10}$. Studies comparing ambient endotoxin concentrations in multiple size fractions, including size fractions larger than $PM_{10}$, are lacking. Furthermore, knowledge on the distribution of endotoxin over particle size fractions in air emitted from farms is limited. Studies that included multiple size fractions were either small in size and included only one farm type (Kirychuk et al. 2010; Schaeffer et al. 2017) or included many farms of various types but only two size fractions (inhalable and respirable) (Seedorf et al. 1998). Nonetheless, size fractions larger than 10 μm may potentially yield considerable endotoxin concentrations. The impact of this on residential level is yet unknown. Dispersion of particles larger than 10 μm is generally limited to shorter distances, but under specific conditions (including high wind speed), dispersion may be more extensive. In our study area, the vast majority of people did not live within a few hundreds of meters of farms. Furthermore, particles larger than $PM_{10}$ are not small enough to penetrate the tracheobronchial and alveolar regions of the respiratory tract. In general, deeper penetration of an air pollution component is accompanied with increased health effects, but whether this also applies for endotoxin is not yet known. We decided to focus on the $PM_{10}$ fraction in the current study; however, obtaining more insight into levels of endotoxin in different size fractions, including particles larger than 10 μm, in ambient air is of the utmost importance.

The supervised stepwise selection approach used selects the model with the highest adjusted $R^2$. However, differences in adjusted $R^2$ may be small between models including a slightly different set of correlated predictor variables. Thus, there is no single absolute LUR model to explain spatial variation of a component in a specific area. Hence, predictors not included in the model are not necessarily unassociated with ambient endotoxin.

No temporal livestock data was available. Information available included animal species and numbers based on licensed data, thus numbers of animals that were actually present at the time of measurement may have differed. Furthermore, no data on manure management was available, which may have an environmental impact as well. Management of manure at the farm has been described to affect farmers' personal inhalable dust and endotoxin exposure (Basinas et al. 2014). In Netherlands, manure storage at the farm can be in reservoirs under- or aboveground. Manure is stored until application or treatment. Land application of manure is allowed solely during a specific period of the year and depends on soil type, land use, and manure type. Manure application is mostly done via injection into the soil.

### Recommendations

To our knowledge, LUR modeling as described in this study has not been described earlier for endotoxin. Relevant insights were gained of use worldwide, although modeling results are not directly applicable to other countries. Studies on transferability of LUR models showed that models for air pollution in urban settings are best developed locally as performance is less when applied to other areas (Hoek et al. 2008; Jerrett et al. 2005; Ryan and LeMasters 2007). Substantial differences in farming practices worldwide enhance transferability limitations of LUR models developed in this study. Yet, knowledge obtained in the current study gives guidance for the optimization of similar campaigns in other parts of the world, which would then allow for comparisons. Not only does this apply for implementation of LUR modeling for endotoxin, but this can also be of use for other bioaerosols. The model evaluation results suggest that LUR modeling of bioaerosol concentrations in relation to livestock requires a larger set of sites compared with urban air pollution studies. Bioaerosol concentrations in general are known to be more variable than inert or chemical exposures; therefore, more measurements are needed for accurate exposure assessment (Spaan et al. 2008b). Additionally, in the case of highly variable and heterogeneous sources like the livestock farming industry, more sampling sites should be included to capture this well. To diminish the impact of very local-scale variation, sampling in parallel at all sites should be considered.

Detailed geographical data on sources, in this case livestock farm characteristics, is important for LUR modeling. Availability of temporal livestock data and alternative study designs would add to modeling possibilities. The design of the current measurement campaign was aimed at developing spatial LUR models. Multiple long-term measurements were performed, one measurement per season, in order to obtain annual average concentrations. For development of spatiotemporal models, besides availability of temporal livestock data and local meteorological conditions, consecutively short-term measurements are recommended. Then insight can be gained into temporal effects, as meteorological conditions, and likely also livestock emissions, can vary substantially within short periods of time. Obtaining multiple samples per

season is recommended for development of seasonal models. Integration of dispersion models into the LUR framework is another interesting option. This approach has been recently applied for $NO_2$ and $PM_{2.5}$ by de Hoogh et al. (2016). However, more information on endotoxin emissions from different animal species is needed before this approach can be applied for the modeling of endotoxin in relation to livestock.

## Conclusion

Substantial spatial variation of annual average endotoxin concentrations was found for 61 residential sites in a livestock-dense area based on measurements of aerosolized endotoxin in the $PM_{10}$ fraction. $PM_{10}$ mass concentration did not appear to be a good surrogate measure of livestock-related ambient endotoxin, as spatial variation of $PM_{10}$ was limited, and did not reflect the spatial variation in endotoxin concentrations. The LUR approach used was successful in explaining variation of endotoxin concentrations between sites on the basis of livestock-related characteristics of the surroundings. LUR models based on a combination of animal-specific livestock-related characteristics performed markedly better than crude proxies. Based on these results, application of the developed endotoxin LUR model is promising for exposure estimation at home addresses of health study participants residing within the same livestock-dense area.

## References

Armstrong BG. 1998. Effect of measurement error on epidemiological studies of environmental and occupational exposures. Occup Environ Med 55(10):651–656, PMID: 9930084, https://doi.org/10.1136/oem.55.10.651.

Babyak MA. 2004. What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models. Psychosom Med 66(3):411–421, PMID: 15184705, https://doi.org/10.1097/01.psy.0000127692.23278.a9.

Basagaña X, Rivera M, Aguilera I, Agis D, Bouso L, Elosua R. 2012. Effect of the number of measurement sites on land use regression models in estimating local air pollution. Atmos Environ 54:634–642, https://doi.org/10.1016/j.atmosenv.2012.01.064.

Basinas I, Sigsgaard T, Erlandsen M, Andersen NT, Takai H, Heederik D, et al. 2014. Exposure-affecting factors of dairy farmers' exposure to inhalable dust and endotoxin. Ann Occup Hyg 58(6):707–723, PMID: 24748620, https://doi.org/10.1093/annhyg/meu024.

Basinas I, Sigsgaard T, Kromhout H, Heederik D, Wouters IM, Schlünssen V. 2015. A comprehensive review of levels and determinants of personal exposure to dust and endotoxin in livestock farming. J Expo Sci Environ Epidemiol 25(2):123–137, PMID: 24280684, https://doi.org/10.1038/jes.2013.83.

Beelen R, Hoek G, Vienneau D, Eeftens M, Dimakopoulou K, Pedeli X, et al. 2013. Development of $NO_2$ and $NO_x$ land use regression models for estimating air pollution exposure in 36 study areas in Europe – the ESCAPE project. Atmos Environ 72:10–23, https://doi.org/10.1016/j.atmosenv.2013.02.037.

Borlée F, Yzermans CJ, Krop E, Aalders B, Rooijackers J, Zock JP, et al. 2017. Spirometry, questionnaire and electronic medical record based COPD in a population survey: comparing prevalence, level of agreement and associations with potential risk factors. PLoS One 12(3):e0171494, PMID: 28273094, https://doi.org/10.1371/journal.pone.0171494.

Borlée F, Yzermans CJ, Van Dijk CE, Heederik D, Smit LAM. 2015. Increased respiratory symptoms in COPD patients living in the vicinity of livestock farms. Eur Respir J 46(6):1605–1614, PMID: 26250492, https://doi.org/10.1183/13993003.00265-2015.

Cambra-López M, Aarnink AJA, Zhao Y, Calvet S, Torres AG. 2010. Airborne particulate matter from livestock production systems: a review of an air pollution problem. Environ Pollut 158(1):1–17, PMID: 19656601, https://doi.org/10.1016/j.envpol.2009.07.011.

Cantuaria ML, Suh H, Løfstrøm P, Blanes-Vidal V. 2016. Characterization of exposure in epidemiological studies on air pollution from biodegradable wastes: misclassification and comparison of exposure assessment strategies. Int J Hyg Environ Health 219(8):770–779, PMID: 27692572, https://doi.org/10.1016/j.ijheh.2016.08.006.

Centraal Bureau voor de Statistiek. 2017a. Landbouw; economische omvang naar omvangsklasse, hoofdbedrijfstype, region [in Dutch]. http://statline.cbs.nl/StatWeb/publication/?VW=T&DM=SLNL&PA=80786NED&D1=0,2&D2=0&D3=a&D4=0,5-16&D5=a&HD=111129-0915&HDR=T,G1,G4&STB=G2,G3 [accessed 17 May 2017].

Centraal Bureau voor de Statistiek. 2017b. Landbouw; gewassen, dieren en grondgebruik naar gemeente [in Dutch]. http://statline.cbs.nl/Statweb/publication/?DM=SLNL&PA=80781ned&D1=0,2,5,12-13,25-26,34,38,40,44,48-49,53-54,59,68,73-76,79-83,87,95,102,106,114,120,124,135,139&D2=0,13,151&D3=0,5,10,14-16&HDR=G1,G2&STB=T&VW=T [accessed 17 May 2017].

Chen Q, Hildemann LM. 2009. The effects of human activities on exposure to particulate matter and bioaerosols in residential homes. Environ Sci Technol 43(13):4641–4646, PMID: 19673245, https://doi.org/10.1021/es802296j.

Clark S, Rylander R, Larsson L. 1983. Airborne bacteria, endotoxin and fungi in dust in poultry and swine confinement buildings. Am Ind Hyg Assoc J 44(7):537–541, PMID: 6613856, https://doi.org/10.1080/15298668391405265.

de Hoogh K, Gulliver J, van Donkelaar A, Martin RV, Marshall JD, Bechle MJ, et al. 2016. Development of West-European $PM_{2.5}$ and $NO_2$ land use regression models incorporating satellite-derived and chemical transport modelling data. Environ Res 151:1–10, PMID: 27447442, https://doi.org/10.1016/j.envres.2016.07.005.

de Hoogh K, Wang M, Adam M, Badaloni C, Beelen R, Birk M, et al. 2013. Development of land use regression models for particle composition in twenty study areas in Europe. Environ Sci Technol 47(11):5778–5786, PMID: 23651082, https://doi.org/10.1021/es400156t.

de Rooij MMT, Heederik DJJ, Borlée F, Hoek G, Wouters IM. 2017. Spatial and temporal variation in endotoxin and $PM_{10}$ concentrations in ambient air in a livestock dense area. Environ Res 153:161–170, PMID: 27984760, https://doi.org/10.1016/j.envres.2016.12.004.

Eeftens M, Beelen R, De Hoogh K, Bellander T, Cesaroni G, Cirach M, et al. 2012a. Development of land use regression models for $PM_{2.5}$, $PM_{2.5}$ absorbance, $PM_{10}$ and $PM_{coarse}$ in 20 European study areas; Results of the ESCAPE project. Environ Sci Technol 46(20):11195–11205, PMID: 22963366, https://doi.org/10.1021/es301948k.

Eeftens M, Tsai MY, Ampe C, Anwander B, Beelen R, Bellander T, et al. 2012b. Spatial variation of $PM_{2.5}$, $PM_{10}$, $PM_{2.5}$ absorbance and $PM_{coarse}$ concentrations between and within 20 European study areas and the relationship with $NO_2$ – results of the ESCAPE project. Atmos Environ 62:303–317, PMID: 26342569, https://doi.org/10.1016/j.atmosenv.2012.08.038.

Elliott L, Yeatts K, Loomis D. 2004. Ecological associations between asthma prevalence and potential exposure to farming. Eur Respir J 24(6):938–941, PMID: 15572535, https://doi.org/10.1183/09031936.04.00006404.

Hamon L, Andrès Y, Dumont E. 2012. Aerial pollutants in swine buildings: a review of their characterization and methods to reduce them. Environ Sci Technol 46(22):12287–12301, PMID: 23127202, https://doi.org/10.1021/es3025758.

Heederik D, Ijzermans C. 2011. Mogelijke effecten van intensieve-veehouderij op de gezondheid, [Potential Effects of Intensive Livestock Farming on Neighboring Residents' Health], Utrecht, Netherlands:Netherlands Institute for Health Services Research [in Dutch].

Heederik D, Sigsgaard T, Thorne PS, Kline JN, Avery R, Bønløkke JH, et al. 2007. Health effects of airborne exposures from concentrated animal feeding operations. Environ Health Perspect 115(2):298–302, PMID: 17384782, https://doi.org/10.1289/ehp.8835.

Hjort J, Hugg TT, Antikainen H, Rusanen J, Sofiev M, Kukkonen J, et al. 2016. Fine-scale exposure to allergenic pollen in the urban environment: evaluation of land use regression approach. Environ Health Perspect 124(5):619–626, PMID: 26452296, https://doi.org/10.1289/ehp.1509761.

Hoek G, Beelen R, de Hoogh K, Vienneau D, Gulliver J, Fischer P, et al. 2008. A review of land-use regression models to assess spatial variation of outdoor air pollution. Atmos Environ 42(33):7561–7578, https://doi.org/10.1016/j.atmosenv.2008.05.057.

Jedynska A, Hoek G, Wang M, Eeftens M, Cyrys J, Keuken M, et al. 2014. Development of land use regression models for elemental, organic carbon, PAH, and hopanes/steranes in 10 ESCAPE/TRANSPHORM European study areas. Environ Sci Technol 48(24):14435–14444, PMID: 25317817, https://doi.org/10.1021/es502568z.

Jerrett M, Arain A, Kanaroglou P, Beckerman B, Potoglou D, Sahsuvaroglu T, et al. 2005. A review and evaluation of intraurban air pollution exposure models. J Expo Anal Environ Epidemiol 15(2):185–204, PMID: 15292906, https://doi.org/10.1038/sj.jea.7500388.

Jonges M, van Leuken J, Wouters I, Koch G, Meijer A, Koopmans M, et al. 2015. Wind-mediated spread of low-pathogenic avian influenza virus into the environment during outbreaks at commercial poultry farms. PLoS One 10(5): e0125401, PMID: 25946115, https://doi.org/10.1371/journal.pone.0125401.

Kirychuk SP, Reynolds SJ, Koehncke NK, Lawson J, Willson P, Senthilselvan A, et al. 2010. Endotoxin and dust at respirable and nonrespirable particle sizes are not consistent between cage-and floor-housed poultry operations. Ann Occup Hyg 54(7):824–832, PMID: 20538718, https://doi.org/10.1093/annhyg/meq047.

Liebers V, Brüning T, Raulf-Heimsoth M. 2006. Occupational endotoxin-exposure and possible health effects on humans. Am J Ind Med 49(6):474–491, PMID: 16586405, https://doi.org/10.1002/ajim.20310.

Liebers V, Raulf-Heimsoth M, Brüning T. 2008. Health effects due to endotoxin inhalation (review). Arch Toxicol 82(4):203–210, PMID: 18322674, https://doi.org/10.1007/s00204-008-0290-1.

Mirabelli MC, Wing S, Marshall SW, Wilcosky TC. 2006. Asthma symptoms among adolescents who attend public schools that are located near confined swine feeding operations. Pediatrics 118(1):e66–e75, PMID: 16818539, https://doi.org/10.1542/peds.2005-2812.

Morgenstern V, Bischof W, Koch A, Heinrich J. 2006. Measurements of endotoxin on ambient loaded PM filters after long-term storage. Sci Total Environ 370(2-3):574–579, PMID: 17005241, https://doi.org/10.1016/j.scitotenv.2006.08.027.

Mueller-Anneling L, Avol E, Peters JM, Thorne PS. 2004. Ambient endotoxin concentrations in PM10 from Southern California. Environ Health Perspect 112(5):583–588, PMID: 15064165, https://doi.org/10.1289/ehp.6552.

Ogink NWM, de Ruiter MM, Groot Koerkamp PWG, van' t Klooster CE. 1997. Emissions of dust and endotoxins from swine house buildings. Pap Am Soc Agric Eng 1997:974093.

Pavilonis BT, Anthony TR, O'Shaughnessy PT, Humann MJ, Merchant JA, Moore G, et al. 2013a. Indoor and outdoor particulate matter and endotoxin concentrations in an intensely agricultural county. J Expo Sci Environ Epidemiol 23(3):299–305, PMID: 23321860, https://doi.org/10.1038/jes.2012.123.

Pavilonis BT, Sanderson WT, Merchant JA. 2013b. Relative exposure to swine animal feeding operations and childhood asthma prevalence in an agricultural cohort. Environ Res 122:74–80, PMID: 23332647, https://doi.org/10.1016/j.envres.2012.12.008.

Pillai SD, Ricke SC. 2002. Bioaerosols from municipal and animal wastes: background and contemporary issues. Can J Microbiol 48(8):681–696, PMID: 12381025, https://doi.org/10.1139/w02-070.

Radon K, Schulze A, Ehrenstein V, van Strien RT, Praml G, Nowak D. 2007. Environmental exposure to confined animal feeding operations and respiratory health of neighboring residents. Epidemiology 18(3):300–308, PMID: 17435437, https://doi.org/10.1097/01.ede.0000259896.62137.84.

Ryan PH, LeMasters GK. 2007. A review of land-use regression models for characterizing intraurban air pollution exposure. Inhal Toxicol 19(suppl 1):127–133, PMID: 17886060, https://doi.org/10.1080/08958370701495998.

Samadi S, Wouters IM, Heederik DJ. 2013. A review of bio-aerosol exposures and associated health effects in veterinary practice. Ann Agric Environ Med 20(2):206–221. [pii], PMID: 23772565, https://doi.org/1052320.

Schaeffer JW, Reynolds S, Magzamen S, Vandyke A, Gottel NR, Gilbert JA, et al. 2017. Size, composition, and source profiles of inhalable bioaerosols from Colorado dairies. Environ Sci Technol 51(11):6430–6440, PMID: 28492313, https://doi.org/10.1021/acs.est.7b00882.

Schiffman SS, Studwell CE, Landerman LR, Berman K, Sundy JS. 2005. Symptomatic effects of exposure to diluted air sampled from a swine confinement atmosphere on healthy human subjects. Environ Health Perspect 113(5):567–576, PMID: 15866765, https://doi.org/10.1289/ehp.6814.

Schinasi L, Horton RA, Guidry VT, Wing S, Marshall SW, Morland KB. 2011. Air pollution, lung function, and physical symptoms in communities near concentrated swine feeding operations. Epidemiology 22(2):208–215, PMID: 21228696, https://doi.org/10.1097/EDE.0b013e3182093c8b.

Schulze A, Rommelt H, Ehrenstein V, van Strien R, Praml G, Küchenhoff H, et al. 2011. Effects on pulmonary health of neighboring residents of concentrated animal feeding operations: exposure assessed using optimized estimation technique. Arch Environ Occup Health 66(3):146–154, PMID: 21864103, https://doi.org/10.1080/19338244.2010.539635.

Schulze A, van Strien R, Ehrenstein V, Schierl R, Küchenhoff H, Radon K. 2006. Ambient endotoxin level in an area with intensive livestock production. Ann Agric Environ Med 13:87–91, PMID: 16841878.

Seedorf J, Hartung J, Schröder M, Linkert KH, Phillips VR, Holden MR, et al. 1998. Concentrations and Emissions of Airborne Endotoxins and Microorganisms in Livestock Buildings in Northern Europe. J Agric Eng Res 70(1):97–109, https://doi.org/10.1006/jaer.1997.0281.

Sigurdarson ST, Kline JN. 2006. School proximity to concentrated animal feeding operations and prevalence of asthma in students. Chest 129(6):1486–1491, PMID: 16778265, https://doi.org/10.1378/chest.129.6.1486.

Smit LAM, Boender GJ, De Steenhuijsen Piters WAA, Hagenaars TJ, Huijskens EGW, Rossen JWA, et al. 2017. Increased risk of pneumonia in residents living near poultry farms: does the upper respiratory tract microbiota play a role? Pneumonia (Nathan) 9:3, PMID: 28702305, https://doi.org/10.1186/s41479-017-0027-0.

Smit LAM, Hooiveld M, van der Sman-de Beer F, Opstal-van Winden AWJ, Beekhuizen J, Wouters IM, et al. 2014. Air pollution from livestock farms, and asthma, allergic rhinitis and COPD among neighbouring residents. Occup Environ Med 71(2):134–140, PMID: 24142990, https://doi.org/10.1136/oemed-2013-101485.

Soukup JM, Becker S. 2001. Human alveolar macrophage responses to air pollution particulates are associated with insoluble components of coarse material, including particulate endotoxin. Toxicol Appl Pharmacol 171(1):20–26, PMID: 11181108, https://doi.org/10.1006/taap.2000.9096.

Spaan S, Doekes G, Heederik D, Thorne PS, Wouters IM. 2008a. Effect of extraction and assay media on analysis of airborne endotoxin. Appl Environ Microbiol 74(12):3804–3811, PMID: 18441112, https://doi.org/10.1128/AEM.02537-07.

Spaan S, Schinkel J, Wouters IM, Preller L, Tielemans E, Nij ET, et al. 2008b. Variability in endotoxin exposure levels and consequences for exposure assessment. Ann Occup Hyg 52(5):303–316, PMID: 18515847, https://doi.org/10.1093/annhyg/men024.

Tager IB, Lurmann FW, Haight T, Alcorn S, Penfold B, Katharine Hammond S. 2010. Temporal and spatial patterns of ambient endotoxin concentrations in Fresno, California. Environ Health Perspect 118(10):1490–1496, PMID: 20494854, https://doi.org/10.1289/ehp.0901602.

Thorne PS, Ansley AC, Perry SS. 2009. Concentrations of bioaerosols, odors, and hydrogen sulfide inside and downwind from two types of swine livestock operations. J Occup Environ Hyg 6(4):211–220, PMID: 19177273, https://doi.org/10.1080/15459620902729184.

Wang M, Beelen R, Eeftens M, Meliefste K, Hoek G, Brunekreef B. 2012. Systematic evaluation of land use regression models for NO$_2$. Environ Sci Technol 46(8):4481–4489, PMID: 22435498, https://doi.org/10.1021/es204183v.

Wing S, Wolf S. 2000. Intensive livestock operations, health, and quality of life among eastern North Carolina residents. Environ Health Perspect 108(3):233–238, PMID: 10706529, https://doi.org/10.1289/ehp.00108233.

Winkel A, Mosquera J, Groot Koerkamp PW, Ogink NW, Aarnink AJ. 2015. Emissions of particulate matter from animal houses in the Netherlands. Atmos Environ 111:202–212, https://doi.org/10.1016/j.atmosenv.2015.03.047.